

Inventors: Yunnan Wu
Anthony Vetro
Huifang Sun
Sun-Yuan Kung

End-to-End Traffic Management and Adaptive Multi-Hop Multimedia Transmission

5 Field of the Invention

The present invention relates generally to the field of network communications, and more particularly to traffic management and content adaptation for multimedia content in a heterogeneous network.

10 Background of the Invention

It is a challenge to transmit multimedia content, from a sender end system to a receiver end system, over heterogeneous networks such as the Internet. The TCP/IP protocol stack in the Internet includes a physical layer, a link layer, a network layer, a transport layer, and an application layer. The lower layers, i.e. network layer and below, provide simple per-packet operations to deliver an addressed packet from the sender to the receiver at best effort, while high-layers, i.e. transport layer and above, provide congestion control by appropriate
20 coordination between the sender and the receiver end systems.

This service model poses major problems with respect to transmitting multimedia. First, the core of the Internet is stateless by design, and thus inadequate for guaranteed quality-of service (QoS). Second, it is difficult to design software for
25 the end systems that can adapt multimedia to the dynamic characteristics of the network, in part due to its heterogeneity.

Adaptive transmission strategies have been developed in two aspects: determining *how* to transmit data, and determining *what* to transmit to best utilize the network resource. These aspects are referred to as traffic management and content
5 adaptation, respectively. In the case of best-effort networks, such as today's Internet, traffic management is concerned with congestion control, i.e., to avoid injecting more traffic than the network can handle, and also to guarantee the fairness among the competing flows in sharing the network resources.

10 Traffic management is essential for the “health” of the network, and the efficiency of the connection. While stability is essential, abrupt actions can adversely affect multimedia traffic. While the AIMD (Additive Increase, Multiplicative Decrease) window control of TCP works well for reliable data transport, the inherent “saw-tooth” window evolution, namely the oscillation of the sending rate, makes AIMD
15 unsuitable for multimedia streaming data. Particularly, when the data is a “live” video where future bit rate requirements are unknown. Recently, so called TCP-friendly solutions use steady state TCP throughput models that consider packet loss ratio and round trip time, to adjust the sending rate. However, a steady state model is less than optimal for regulating transient behaviors of streaming data,
20 such as a “live” video feed.

Congestion avoidance methods can predict and control the network state to prevent congestion in the first place. For example, L. Brakmo and L. Peterson, “*TCP Vegas: End-to-End Congestion Avoidance on a Global Internet*”, IEEE
25 Journal on Select Areas of Communications, vol. 13, No. 8, pages 1465-1480, Oct. 1995, use a heuristic window-based process that adjusts the transmission window

linearly when a difference of expected rate and an actual rate is greater than an upper threshold or less than a lower threshold.

H. Kanakia, P. P. Mishra, and A. Reibman, “*An adaptive congestion control scheme for real-time packet video transport*”, Proc. ACM SIGCOMM, Sept. 1993, determine the bits for every video frame by requiring that intermediate routers feed back queue size information. They predict a bottleneck router queue size evolution over time, and try to keep the bottleneck queue size at a constant level. Their system is effective for adaptive congestion control in a best-effort network, and achieves graceful video quality degradation during network congestion.

However, there are some drawbacks with that method. First, they assume that each router monitors the queue occupancy and the service rate per connection. This information is fed back in response to periodic query packets. Such an assumption is not realistic in today’s stateless Internet core. Second, the method directly works with bit allocation for a video frame from feedback on network conditions. That is, the method does not separate the roles of traffic management and content adaptation. Third, the method does not consider packet losses.

Due to the heterogeneity of the network, prior art end-to-end adaptive transmission do not capture network dynamics very well. For example, it is well known that the Internet’s Transmission Control Protocol (TCP) suffers from serious performance degradation when running over a path containing a wireless link.

As a partial remedy, proxy servers and filters can be placed in the network. For multimedia multicasts, filters are used to construct a distribution tree to make efficient use of the network resources, and to accommodate end systems with different link bandwidths. When the sending end system is too busy, filters are used in place of the sender to adapt the content during network congestion.

Caching and pre-fetching with proxy servers can take advantage of similar access among many end systems. Application level filters can be used as incrementally deployable alternatives to programmable routers for placing user-defined computations into the network. Real-time multimedia transcoding is a specific example.

Efficient collaboration between the filters and the end systems remains a problem. Lack of collaboration can result in inefficiencies, duplication, and negative interaction. Therefore, it is desired to provide a system that coordinates the filters and end systems efficiently in terms of both performance and system resource, especially when dealing with multimedia content in a heterogeneous network.

Summary of the Invention

The present invention provides an end-to-end traffic management system and method for multimedia content delivery over best-effort networks. A multi-hop video communication system is also described. Traffic management and content adaptation are done collaboratively by end systems and relays in the network. Multi-hop management achieves higher throughput without increasing the risk of overloading the network, and increasing end-to-end delays. The relays according

to the invention include an error withdrawal feature so that abrupt decreases in quality of the content are avoided.

More particularly, the invention provides a method for managing traffic over a
5 channel of a network connecting a sender end system and a receiver end system.
The traffic includes multimedia packets. The channel is modeled as a queue
having an associated queue occupancy. A time series of samples for a service time
experienced by each packet that is sent is updated based on the times when packets
are sent and the times when feedback messages are received. A most recent queue
10 occupancy is then predicted based on the time series, and the next packet is sent
according to the predicted queue occupancy.

Brief Description of the Drawings

Figure 1 is a block diagram of a communications system that uses end-to-end
traffic management according to the invention;

Figure 2 is a block diagram of a queue model used for end-to-end traffic
management according to the invention;

Figure 3 is a graph of cumulative send and feedback packets, and queue
occupancy;

Figure 4 is a graph of cumulative send packets, throughput at playback times;

Figure 5 is a flow diagram of a process used by the method according to the invention;

Figure 6 is a graph of over-allocation and bandwidth reduction;

5

Figure 7 is a flow diagram of end-to-end traffic management according to the invention; and

10

Figure 8 is a block diagram of a queue model for multi-hop traffic management according to the invention;

10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180
2181
2182
2183
2184
2185
2186
2187
2188
2189
2190
2191
2192
2193

implemented at independent high-level layers of the network protocol stack, as described in greater detail below, i.e., above the network layer.

The communications channel 104a-b is shown in a highly simplified form. In practice, the channel passes through many intermediate switches, firewalls, gateways, routers. At the physical layer, the channel can include, wire, wireless, RF, infra-red, micro-wave wireless and satellite links, co-axial and optical cables, each having their own bandwidth characteristic. The network can include the world-wide telephone system, the Internet, the World-Wide-Web, broadband, baseband, broadcast, satellite, and multicast networks.

In one embodiment, a multi-hop version of the system 100 can include one or relays 103, described in greater detail below.

Feedback Messages

In response to receiving packets, the receiver 102 sends feedback messages 106-107 upstream, with respect to the flow of the traffic 105 from the sender end system 101 to the receiver end system 102. Each feedback message 120 comprises two parts: application feedback data 121 and transport feedback data 122. These data are suitable for processing by the content adaptation layers and the traffic management layers 111-112, respectively. Specifics of the feedback messages are described in greater detail below.

End-to-end traffic management 112 according to the invention includes two phases: a learning phase and a near steady phase. During the learning phase, data

are transmitted at a relative slow and “safe” constant rate to collect a history of feedback messages. Essentially, the network characteristics are “learned” or determined without overloading the system. After the learning phase, the data rate is adapted to best utilize the available bandwidth without congestion in the near steady phase.

As shown in Figure 2, the end-to-end channel 104a-b can be modeled as a FIFO queue Q 200 and delay elements τ_1, τ_2, τ_3 . Of course in a real network, the channel contains an unknown number of routers, switches and data links with various unknown bandwidths and traffic patterns and delays.

The total amount of delay, aside from a time-varying packet service time at the queue, is $\tau = (\tau_1 + \tau_2 + \tau_3)$. Due to random shared access by many competing connections, the packet service time at the queue on the channel 104 is modeled by a random time series $s(n)$ 210, where n represents a packet number. The time series 210 is locally stationary.

Figure 3 models the traffic management at the sender 101. The x-axis 301 indicates time, and the y-axis 302 the cumulative number of packets sent. At a current time 303, the sender’s process maintains a cumulative packet sending curve 311, i.e., the total number of packets sent at any point in time, and a cumulative feedback curve 312. The two curves essentially “count” the total of number of packets sent and received on the channel over time. In order to determine the input and output rate at the queue 200, the cumulative feedback curve 312 is shifted back the delay time $\tau_1 + \tau_2 + \tau_3$ 305 to a position 313.

This divides the time axis 301 into three time intervals 321-323. Time interval 321 corresponds to packets sent and acknowledged by the receiver 102, interval 322 corresponds to packets sent but not acknowledged, and interval 323 corresponds to unsent packets. At a current time 303, the vertical difference 306 between the cumulative sending curve 311 and the shifted cumulative feedback curve 313 models queue occupancy. The queue occupancy reflects the amount space in the queue is used, e.g., an occupancy of zero could mean the queue is empty and an occupancy of one means the queue is full. For a particular packet, the horizontal difference 305 between the two curves 311-313 models the time spent in the queue, i.e., the service time.

In order to estimate a most recent queue occupancy and to regulate the transmission of packets in the future, the time series $s(n)$ 210 is predicted some period ahead in time. For longer prediction periods, local details become less important.

Therefore, the traffic management module 112 according to the invention uses a multi-timescale linear prediction method 700 as shown in Figure 7. The process 700 is responsive to feedback messages 701. When a feedback message is received, the cumulative sending curve 311 and the cumulative feedback curve 312 are updated 710 to reflect an advance in time. For the interval 321, a time sample t_s is added 720 to the time series $s(n)$ 210 for each packet sent, where: $t_s = \text{departure time of packet } n - \max(\text{departure time of packet } n-1, \text{arrival time of packet } n)$.

The departure time and the arrival time are both with respect to the queue 200.

The multi-timescale linear prediction 730 predicts the time series $s(n)$ some steps ahead. This is done by subtracting the mean μ for the observed time series $s(n)$ from each $s(n)$ to produce a zero-mean time series. The zero-mean time series is then passed to the zero-mean multi-timescale linear prediction. After the prediction, the mean is added back to get the prediction output, and the method terminates 750 until a next feedback message is received.

The zero-mean multi-timescale linear prediction 730 can be performed by known decimation, linear prediction and interpolation in a series with a decimation factor of $2k$ in a k -th timescale. In this way, at any timescale, two steps ahead can be predicted. The taps of a prediction finite impulse response (FIR) filter can be obtained by the Yule-Walker equation, or using an adaptive filtering algorithm such as LMS, see S. Haykin, "*Adaptive Filter Theory*," Third Edition, Prentice-Hall, 1996.

The prediction can then be used to update the packet transmission schedule 760 for a small number of next packets. With the observed and predicted service time series $s(n)$, the departure times from the queue can be determined for all the packets that have been sent. With respect to the model of Figure 4, this means that the shifted cumulative feedback curve 313 can be extended. The estimated queue occupancy at the boundaries of the intervals 321-323 is represented by the vertical distance 306 between the sending curve 311 and the sending feedback curve 313, at the current time 303.

Transmission 770 of the next few packets is then regulated to gradually drive the queue occupancy to the desired queue occupancy. A simple analysis with the M/M/1 queue model suggests setting the queue length to be one, see M. Schwartz, *“Broadband integrated networks,”* Prentice-Hall, 1996. An M/M/1 queue has a
 5 Poison arrival time, and an exponential service time, and a single server.

If the estimated queue occupancy is less than one ($Q < 1$), then one packet is transmitted immediately, and another packet is transmitted whenever a feedback message from a previous packet is received. Otherwise, if the estimated queue occupancy is one ($Q = 1$), transmit a next packet whenever a feedback message for a previous sent packet is received. And otherwise, if the estimated queue occupancy is greater than one ($Q > 1$), transmit the next packet until the queue occupancy again becomes one.

Although this prediction process avoids most congestion, packet loss can also happen due to errors in the prediction. Because a packet is dropped when there is no room in the queue Q 200, congestion loss signals the size of the available queue occupancy for the channel. For example, if it determined that packet #2 is lost when the feedback of packet #3 arrives, then the available queue occupancy
 20 can be updated.

Similarly, when a packet gets through, another sample of the available queue occupancy is collected. Therefore, instead of going directly from step 710 to step 720, an alternative path through step 715 identifies those packets lost since the last
 25 feedback message based on sequence number information in the feedback message.

Because a packet is dropped when there is no room in the queue, packet loss signals the size of the available queue occupancy for the connection. For every packet lost, a “trouble-making” queue occupancy 601 curve, see Figure 6, can be determined from the sending and shifted feedback curves. Then, update the available queue occupancy according to the following equation:

$$new_q_occupancy = \min(old_q_occupancy, \text{“trouble-making” queue occupancy} - 1).$$

Similarly, when a packet arrives successfully, collect another sample of the available queue occupancy. For every successfully sent packet, the “healthy” queue occupancy 306 is inferred, and the available queue occupancy is updated according to:

$$new_q_occupancy = \max(old_q_occupancy, \text{“healthy” queue occupancy}).$$

The available queue occupancy information can also be used in estimating the queue occupancy for interval 322, this is, packets sent but not acknowledged. This leads to a modified queue occupancy estimation step. If the queue occupancy is too large when a packet arrives, then a packet loss can be predicted. The modified queue occupancy estimation continues assuming the packet is discarded. Such an early loss warning can also be useful for a video encoder. The encoder can avoid using that packet for a reference frame to reduce error propagation. The available queue occupancy can also be used to control the speed the queue occupancy is driven to the desired value. If the available queue occupancy is large, the speed of control can be slow, and vice versa.

So far, the traffic management method 700 assumes a steady inverse flow of feedback messages. However, the feedback can be compressed to reduce the amount of control traffic. Because the sender is doing prediction, a feedback message can be avoided when a new sample of the time series is predictable. In this case, the prediction of the sender can be used instead.

Multi-Hop Video Transmission System

Figure 1 also shows that one or more relays 103 can be inserted between the end systems 101-102. In this case, the relays 103 can also process the feedback messages 106-107. At each relay 103, the channel is presented as an input link 104a and an output link 104b. In contrast to the end-to-end traffic management described above, where only the end systems 101-102 perform traffic management and content adaptation, here the relays 103 provide traffic management and content adaptation within the network. As before, traffic management and content adaptation modules 131-132 are implemented at independent high-level layers of the network protocol stack. The relay 103 also includes a buffer 150 for temporarily holding received data and feedback messages.

The one or more relays 103 can perform as: observer, advisor, and controller. A particular relay can perform one, or a combination of these functions. As an observer, the relay collects QoS information for the input and output links to which it is directly connected, and the traffic flow it is monitoring. As an advisor, the relay provides content adaptation predictions and suggestions based on the QoS information collected by observer relays. For example, packet coloring can be used to assess priority relations. As a controller, the relay performs traffic

management and content adaptation, taking into consideration the resource trade-off of multiple competing connections. In other words, the relays enable the monitoring and adapting of traffic state inside an otherwise stateless network.

5 Unlike prior art routers, which commonly performs only per-packet processing at the lower layers of the communications protocol, the relays 103 according to the invention can identify a traffic flow, that is a sequence or “flow” of related packets, such as a video or audio program. Hence, higher level adaptation and management can be provided. Also, the blocks representing the upper layers 131-10 132 are intentionally shown to be “thinner” in Figure 1 to emphasize that they are considerably less complex than the comparable layers 111-112 in the end systems 101-102.

Separated Traffic Management and Content Adaptation

The two layered modules 131-132 are implemented as two distinct processes with well defined interfaces. Thus, the layers 131-132 can operate independent of each other. Balakrishnan et al. in “*An Integrated Congestion Management Architecture for Internet Hosts*,” Proc. ACM SIGCOMM, Sept. 1999, describe sharing

20 congestion control among connections with identical source and destination pairs at an end host. Here, the separated design of modules 131-132 at the relay can share the traffic management among peer connections with identical relay and destination, (or another relay) pairs. Sharing of traffic management also means some common feedback packets for network link condition can be reduced.

25 Because the relay is assumed to serve more than one connection, sharing is

predicted to happen more frequently. Thus, it is advantageous and cost-effective to have a common traffic management.

With the multi-hop system 100 according to the invention, the QoS problem in the stateless Internet is addressed by coordinating traffic management and content adaptation not only in the end systems, but also in the relays which now can maintain state of the traffic and content in the otherwise stateless network.

Multi-Hop Traffic Management

In the multi-hop management system 100 there are at least two control loops, one for the input 104a, and another one for the output link 104b. Because these two loops are always shorter than the total end-to-end loop, the relay 103 can better track the network state, and react more quickly to changes in the network state.

In this case, the traffic management includes learning, speed-up, and steady phases. During the learning phase, packets are transmitted slowly to learn the network characteristics. During speed-up phase, each local loop runs at a maximal speed to best utilize the available bandwidth without congestion.

Consider the case of a high-bandwidth Internet channel connected to a downstream low-bandwidth wireless channel. Because of the mismatch in the throughput of the different loops, the traffic is buffered at the relays before the bottleneck builds up. In the long run, the faster loops are constrained by a bottleneck slower channel. Such a constraint can be applied to the traffic management module by estimating a throughput at the receiver 102, and keeping

the total outstanding packets in the system roughly at a constant value. This corresponds to the steady phase.

Figure 8 shows a model 800 of the multi-hop system 100 with multiple loops.

Here, the system can be modeled by multiple queues 801 and 802, one for each loop. In this case, the sender 101 receives multiple feedback flows: one flow 811 from the receiver 102, and one flow 812 from each relay in the end-to-end path. Therefore, the sender 101 performs multiple predictions, one for each feedback flow.

Now, in the packet transmitter 770 of Figure 7, a total cumulative sending at the sender 101 and a total cumulative arriving at the receiver 102 can be considered. Packets are scheduled at first according to the local loop, that is, based on the cumulative sending at the sender and the cumulative feed back from the relay. If these are within the buffer constraints, then the packet transmission is scheduled. Otherwise, transmission of the next packet is deferred by a time unit, and test again.

Multi-Hop Content Adaptation

Given the available bandwidth determined by the traffic management module, the content adaptation module schedules the data to be transmitted. When the available bandwidth is limited, transcoding may be applied to gracefully degrade the quality. Performing one adaptation over another depends on the characteristics of the content, the playback time constraints, and the available bandwidth. This is often formulated as a rate-distortion (R-D) optimization problem.

Hsu et al., in “*Rate control for robust video transmission over burst-error wireless channels*,” IEEE JSAC, vol. 17, no. 5, May. 1999, described optimized video adaptation with playback time constraints for variable bit-error-rate wireless channel. They showed that the delay constraints are equivalent to bit budget constraints from future channel rates.

In the multi-hop management system, both the sender 101 and the relay 103 can perform content adaptation. There is a trade-off as whether the adaptation should be sender-centric or relay-centric.

As shown in Figure 1, the data traffic flow from the sender 101 to the receiver 102 and the feedback messages 106-107 travel the opposite way. Therefore, the sender 102 has better knowledge of the content state while the relay 103 has better knowledge of the traffic state.

When a high bandwidth link is connected to a low bandwidth link at the relay 103, a large buffer 150 can be allocated at the relay 103 to gain content knowledge. An application for this strategy is at the edge of the network. The internal links of the network may operate at rates of 10Mbps or higher. However, the last hop in the network, from a relay, e.g. a cellular base station, to a portable end system, e.g. a cell phone, is via a wireless link operating nominally at 10Kbps. This is a thousand-fold drop in bandwidth. Hence, a relatively large buffer can be allocated in trade-off for the better content knowledge.

Alternative, it is possible for the relay 103 to acquire knowledge about future content by transmitting a content outline before transmitting the actual content. Then, the relay can “learn” the characteristics of the content to be sent ahead of time.

5

However, both solutions consume system resources and more bandwidth than necessary. Consequently, the system 100 prefers a sender-centric adaptation, and use the content adaptation module 131 at the relay 103 to “withdraw” erroneous over-allocation made by the sender 101 when there is a sudden decrease in the available bandwidth. Because the over-allocation withdrawal property is only called for when there is a sudden drop in bandwidth, it has minimal impact on processing delay.

10

Content Adaptation Procedure

A sliding window content adaptation procedure is used for rate-distortion optimized resource allocation with delay constraints. The adaptation procedure is applicable to both the sender 101 and the relay 103. Although prior art sender adaptation can be adapted to work at the relay, there are two additional requirements: the inputs to the relay are on-line data streams that may be subject to packet loss, bit error, delay, and delay jitter, and the relay minimizes latency.

20

As shown in Figure 4, where the current time is T 401 and a start-up delay is D 402, the y-axis indicates the cumulative bits (CB) at the end of each frame of a video, the throughput curve 411 fed back from the receiver 102 maps the playback

25

time constraints into bit budget constraints. Obviously, each frame (F_1, \dots, F_4) must arrive before its scheduled playback time

Because the adaptation refers to future throughput, the throughput curve 411 has to be predicted by the traffic management module. The playback times are shown as vertical dashed lines 430. Equivalently, this means that the cumulative bits (CB#) used must be less than a bound at intersections 450 between the throughput curve 411 and the playback times 430 of the frames.

Figure 5 shows the step of the content adaptation procedure 500 according to the invention. For all the frames in the buffer, i.e., within the current window, step 510 collects the rate-distortion (RD) characteristics. Step 520 partitions the available bandwidth into shares for each frame, while minimizing the received distortion under the bit budget constraints. This optimization problem can be solved with dynamic programming, or the Lagrange multiplier, see Hsu et al. above. If the buffer is full when the frame arrives, step 530 compresses the buffered data if possible, otherwise the frame is discarded. When the frame is transmitted, step 540 slides the window forward to the next frame.

Over-Allocation Withdrawal

The relay 103 can coordinate with the sender to do more than just buffer the content in transit from the sender to the receiver. The sender has full access to the content, and hence, the sender can do long-term prediction for bandwidth allocation. However, because the sender can sometimes incorrectly estimate future available bandwidth due to rapidly changing conditions in the network, bandwidth

may be either over, or under allocated. Over allocation wastes network resources, and under allocation can degrade quality.

For example with reference to Figure 6, which also plots cumulative bits (CB) versus time, frame F3 was over-allocated. However, when the *relay* 103 determines the bandwidth allocation for frame F2, the relay can realize an over-allocation for frame F3 before it is sent out from the relay. This results on a reduced bandwidth allocation for frame F3, as shown with by arrow 500.

Consequently, if the *relay* 103 notifies the sender 101 of the reduced allocation, then the sender can perform allocation assuming some previous over-allocation has been withdrawn. This is equivalent to reducing the current allocation, namely the total bits consumed by all previous frames.

Signaling overhead can be reduced if the sender predicts the reduction behavior at the relay 103. The sender does not need to predict *what* is reduced, instead the sender only predicts *how much data* are to be reduced. Specifically, the number of feedback messages sent by the receiver end system could be reduced when the predicted queue occupancy is within a predetermined error measure.

The sender can check all the frames, from the oldest to the newest, that have been sent but not acknowledged by the receiver 102. If a particular frame exceeds the frame playback time, the sender can assume that the relay 103 will reduce the frame to the frame playback time, and adjusts the total bits consumed.

Although the invention has been described by way of examples of preferred embodiments, it is to be understood that various other adaptations and modifications can be made within the spirit and scope of the invention. Therefore, it is the object of the appended claims to cover all such variations and modifications.

[illegible]